DOCUMENT RESUME

ED 170 333 TM 008 567

AUTHOR Bredderman, Ted
TITLE Elementary School Process Curricula--A Meta-Analysis.
PUB DATE [78]
NOTE 12p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Compensatory Education Programs; *Curriculum Evaluation; Elementary Education; *Elementary School Science; *Evaluation Methods; Evaluation Needs; Experimental Curriculum; *Process Education; *Program Effectiveness; *Program Evaluation; Research Methodology
IDENTIFIERS *Meta Evaluation; *Project Follow Through

ABSTRACT

The effects of the Follow Through Program as determined by meta-analysis techniques were reviewed and summarized. The general conclusion that more direct instructional models had greater positive effects on measures of basic skills in mathematics and reading than less direct, developmental models was supported. However, the validity of the Follow Through Evaluation as a general test of the worth of instructional models was questioned by contrasting the Follow Through results with a similar meta-analysis of elementary school science process programs. The science process programs had substantial effects on the measures which matched their objectives--that is, science process measures--and minimal or negative effects on more distant objectives. It was concluded that testing a broad spectrum of outcomes, including attitude and creativity, is essential in evaluating a curricular instructional model. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ELEMENTARY SCHOOL PROCESS CURRICULA - A META-ANALYSIS

Ted Bredderman

State University of New York at Albany

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY Theodore H Bredderman TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED170333

At the present time those interested in curriculum issues appear to be in a calm after a storm. Evaluation, accounting for past actions and critique seem to be the order of the day. Periods of activity provide the data for periods of reflections, at least for those who take life seriously. Given the spirit of the times, I would like to present reviews of some data gathered in the not too distant time of major curriculum reform and then reflect on their meaning for curriculum evaluation.

The first set of data comes from the widely publicized Follow Through evaluation and the second from a meta-analysis of my own on a collection of evaluation studies of the elementary science process curriculum programs. The two combined analyses encompass over 200 individual studies of implemented curricula.

The Follow Through Evaluation

Follow Through has been called "the largest and most expensive, federal educational experiment in this country's history." (House, Glass, McLean, and Walker, 1978, p. 128) It involved 79,000 children, cost half a billion dollars to operate and nearly 50 million to evaluate. It was an attempt to maintain and build upon any gains, during the primary schools years, which children may have realized earlier in head start programs.

For a variety of reasons a planned variation paradigm was adopted. More than 20 different educational models were tested in over 180 primary grade settings across the country. These models ranged from open to direct instructional approaches, and from stressing so called basic skills to stressing problem solving and affective ends. Outcome instruments used at each site reflected to some extent this same breadth of intentions. Instructional events were also measured through the use of trained classroom observers armed with systems for categorizing classroom settings and events.

For my present purposes it suffices to try to encapsulate the general findings of the Follow Through evaluations which can be gleaned from several sources. These sources include an official evaluation commissioned by the Office of Education (Stebbins, St. Pierre, Proper, Anderson, and Cerva, 1977) and a critical assessment of this official evaluation offered by House, Glass, McLean, and Walker (1978). Also used was a reevaluation using meta-analysis techniques reported by Mary M. Kennedy (1978) of the U.S. Office of Education and a summary of the classroom observation portion of the evaluation reported by Jane Stallings (1975).

Conclusions were reached by combining the results from several

TM 008 567

sites using a particular curriculum model and comparing them with a similar aggregation of results for other models. All of these assessments seem to me to reach the same general conclusion. All leave one with the distinct impression that the more direct structured methods of education are more effective for Follow Through students in improving language and math computation. This impression is even left by the critics who set out to show that it isn't so. The various models line up in roughly the same order under several different forms of analysis with only the direct models showing any substantial advantage. This conclusion holds also for the classroom observation data which showed generally positive correlations with achievement in math and reading and the occurrence of direct teaching events. Conversely, negative correlations between these outcomes and indirect teaching were found. Classroom events which competed with direct teaching for instructional time correlated negatively with achievement. These observational data seem even more damaging to advocates of indirect teaching for these children and these outcomes in that they directly relate classroom events to achievement without resting the case on the assumption that a particular model of instruction had been implemented at a site.

The more indirect instructional models do fair slightly better on measures of problem solving than they do on basic skills. The concern is that they do not do better than the direct models on these more broadly based outcomes. This suggests that in using the direct rather than the indirect approaches nothing is lost in problem solving outcomes and something is gained in basic skills.

So far I have expressed the results of the Follow Through evaluation by comparing instructional models relative to other instructional models. But of course at each site a model was compared with a control group not with another model. Presumably, the control groups underwent fairly traditional approaches to education and these approaches were reasonably comparable from site to site. It is also assumed that the control group students were comparable to those of the Follow Through group which underwent the instructional model treatment. This latter assumption is acknowledged not to be so by all the evaluators, and how this problem should have been treated in the evaluation is a bone of contention among evaluators. I will drop this methodological issue at this point since there is no way of handling it quickly. In doing so I don't mean to diminish its importance.

It will be useful now to take a short diversion into another area of methodology, namely, the problem of aggregating the results of what amount to separate evaluation studies carried out by different investigators at different sites. The traditional approach is to do a more or less qualitative and subjective assessment of each individual study, select those which seem most methodologically sound and then conclude that their results are closest to the truth. A second approach is the one followed by the commissioned evaluation of Follow Through. In this case frequencies of statistically significant differences between experimental and control groups are analyzed. If one model of education shows statistically significant

results favoring it over control groups more frequently than expected by chance, it is concluded that the model is more effective than the control treatments.

The third approach, which has also been applied to the Follow Through data, is to analyze the actual size of the model's effect compared with that of the control group at each site and to treat the measure of relative effect at each site as a data case. This requires the introduction of a measure of effect of a treatment termed "effect size".

As a measure of a treatment's success, effect size has certain advantages which have been presented by Gene Glass in the Educational Researcher (1976). Effect size is calculated by subtracting the mean of the control group from that of the experimental group on any outcome measure and then dividing this difference by the pooled standard deviations of the two groups. If the effect size is one (1) on a particular outcome measure the average student, who naturally would be at the fiftieth percentile in the control group, has moved to the 84th percentile; that is, the student has moved 34 percentile points or one standard deviation, presumably by virtue of the fact that the student is in the experimental group. A negative effect size means that being in the experimental group rather than the control group has caused the average student to perform more poorly on the outcome measure.

Table 1 shows some of the actual data from the Follow Through evaluation for selected models in terms of effect size. For each of several outcomes a mean effect size is given for a number of studies carried out at different sites on a particular model. For the sake of simplicity I have selected two models which have been characterized as cognitive (House et. al. 1978) and one model which has been termed a direct basic skills curriculum. As mentioned before in non-quatitative terms the average student in the directed, basic skills curriculum groups substantially outperformed the average student in the respective control groups, especially on measures of language and math computation. The control groups actually performed better than the experimental groups on the average for the cognitive curricula as indicated by the negative effect sizes.

Elementary School Science Process Curricula

At this point I will describe some preliminary results of another meta-analysis which I have been working on for some time and then try to tie the two meta-analyses together at least on a few points. I hasten to add that the meta-analysis which I am reporting on now is not complete. What is presented here are some preliminary findings.

Three particular elementary science programs ultimately dominated

TABLE 1

Mean Effect Sizes of Selected Follow Through Models on Selected Outcome Measures.[a]

	Education Development Center Model (EDC) Affective/Cognitive[b]	Cognitive Curriculum Model Cognitive/Conceptual[b]	Direct Instruction Model Basic - Skills[b]
Metropolitan Achievement Test			
Language	-.24	-.29	.84
Math Computation	-.16	-.21	.57
Math Problem Solving	-.20	-.09	.17
Reading Comprehension	-.26	-.20	.07
Raven Progressive Matrices	-.12	-.11	-.08

[a] From Kennedy (1978)

[b] House et.al. (1978) characterization of model

the elementary science curriculum reform movement of the 1960's and all three departed from the traditional science curricula of any educational level. The process of science became the central goal. Textbooks were out. Reading as a tool to learning science was played down. Even calculation was reduced in importance. Understanding, reasoning, problem solving, and the investigative processes became the core of the new curricula. Content served process ends rather than the other way around. Concepts introduced were limited to those which made a significant contribution to investigating and reasoning power. Beyond these common departures from tradition the three major projects did present somewhat unique products. One product, the Elementary Science Study (ESS), represents a basically unstructured approach. In this program physical phenomena are often introduced to spark curiosity and give rise to fairly open ended investigation. No particular student conceptualizations are insisted upon. Units are not sequenced. Testing is done informally, if at all. A second product, Science A Process Approach (SAPA), pursues process objectives almost exclusively. The objectives are prespecified and taught in a particular sequence. The processes are taught in explicit terms with ample opportunity to apply them in laboratory activities. A third product, The Science Curriculum Improvement Study (SCIS), is more of a blend of process and concept teaching, although only a few concepts useful in analyzing physical phenomena are taught. Likewise, there is a blend of open-ended and direct teaching. Some lessons allow free exploration; some require the introduction of a concept followed by its application. In summary, ESS and SAPA both concentrate on process almost exclusively. However, ESS uses an open instructional approach; SAPA, a direct one. SCIS is a blend program, blending process and content and open and closed instructional approaches.

All three programs became available for school use in the mid 1960's. Millions of public dollars were spent on their development and millions more spent on implementing the programs and training teachers. At their peak years, the early part of this decade, it was estimated that approximately one third of the nation's elementary children were using at least some portions of one or more of the programs. This, like Follow Through, was a grand and costly national experiment.

In contrast with Follow Through, no orchestrated attempt was made to assess the effectiveness of these process programs. However, many individual evaluation studies were done and reported in literature. After a thorough literature search I managed to identify and acquire the reports of 60 such studies. Only studies which included control groups and which reported means and standard deviations on outcome measures were included. No study which met these two criterion was omitted from the data analysis. There was a great deal of variation among the studies on a number of dimensions. Because of this, each study was examined and the values for 34 study variables were coded for computer analysis. The examination of relationships between study variables and findings

is not complete and will be included in a later report.

The unit of analysis which seemed appropriate for the meta-analysis was a single pairwise comparison of control and experimental group on a particular outcome measure. Each study might make several pairwise comparisons, by comparing the experimental and control groups on several outcome measures or by breaking each group down into subgroups such as first graders and second graders or low ability and high ability students. The total number of pairwise comparisons in the 60 studies amounted to nearly 300.

Outcomes measured were categorized into six broad areas; creativity, affect, language, general cognitive abilities, science content, and science process. Table 2 indicates the mean effect sizes on the 6 outcomes for each of the three process programs. Unfortunately fewer studies have been done on ESS than on the other two programs, yet some tendencies are obvious. The open approach is especially effective on measures of creativity and detrimental on content measures. It shows a moderate advantage, effect size of about .3, on process. SAPA as expected shows a substantial advantage on process, no advantage on content or language, and modest advantages on other measures. The blend program, SCIS, shows a modest to high advantage on all outcomes.

This general pattern suggest three conclusions. First, you get at least what you teach for. All three process programs result in greater achievement on process measures than control group treatments. Second, there is a conservation of total curriculum effect. If one aspect of the curriculum is stressed, one or more other aspects are diminished with regard to achievement. This is suggested by the fact that if a grand mean effect size is calculated for all outcomes for each program, the three grand means are almost the same, but the relative contribution to each outcome measure is different for each program. Third, the more direct approach is superior for convergent outcomes but not for fostering general creativity. This is shown by comparing just the open (ESS) and direct (SAPA) program both of which stress process.

The 50 or 60 evaluators of the process curricula, unlike Follow Through evaluators, used widely different designs and different outcome instruments. These differences and others suggest caution in making comparisons of the results of the two meta-analyses, yet there are some areas in which a reasonable observer would expect to find similar tendencies. Table 3 presents data from the process program meta-analysis for studies involving students which could be classed as disadvantaged based on geography (rural or inner-city), ability, or socio-economic status.

Comparing the data in Table 1 and Table 3 it can be seen that results of the process curricula are most similar to those of the direct instructional model of the Follow Through analysis, at least in direction and range of effect sizes. One reason for

TABLE 2

Mean Effect Sizes of Selected Elementary Science Process Programs for Outcome Categories for Which at Least Six (6) Comparisons Were Available.

	ESS Comparisons Open/Process	SCIS Comparisons Open, Direct/ Process, Content	SAPA Comparisons Direct/Process
Science Process Measures	.33 (12)[a]	.35 (50)	1.01 (38)
General Cognitive Measures		.58 (15)	.24 (16)
Affective Measures		.11 (13)	.16 (7)
Creativity Measures	.63 (14)	.22 (7)	.10 (15)
Science Content Measures	-.24 (12)	.52 (14)	.00 (13)
Language Measures		.08 (7)	-.03 (14)

[a] Numbers in parenthesis indicate the number of comparisons

TABLE 3

Mean Effect Sizes of Combined Elementary Science Process Program Studies on Outcome Categories for Which at Least Six (6) Comparisons Involving Disadvantaged Students Were Available.

	Combined Elementary Science Process Program Comparisons[a]
Science Process Measures	1.02 (30)[b]
General Cognitive Measures	.36 (13)
Affective Measure	.22 (7)
Language Measures	.52 (12)

a Does not include any comparisons involving ESS since fewer than 6 comparisons were available on disadvantaged students for ESS.

b Numbers in parentheses indicate the number of comparisons.

this may be that none of the 65 comparisons of the process curricula for disadvantaged students were made in studies involving ESS, the open process approach. Thus SCIS and SAPA programs which both use, at least in part, fairly direct teaching methods would be expected to be comparable to other direct teaching methods. The scope of outcomes tested is greater in the process studies giving some evidence that curriculum areas other than the basic skills can be affected by planned intervention. The effect sizes indicated for the ESS program in Table 2 point out the need to include outcome measures spanning a broad spectrum of potential program effects. Whether or not the less directed models of the Follow Through experiment might have presented a profile of effects similar to that of ESS; high in process and creativity and low in convergent outcomes, is left unanswered by the Follow Through evaluation.

Some Final Impressions

The two meta analyses which I have described leave behind five major impressions. The first impression is the need for many studies to settle any question of curricular value. The variation among reasonable replication studies is great. The fault for this is not likely to rest, in any large measure, with researchers but is fundamental to the very nature of experimental sampling. Careful evaluation designs may reduce variance among findings but it can not be expected to eliminate it to any major extent.

The second impression is that an evaluation of any curriculum program which does not measure a broad range of potential curriculum effects is bound to be misleading. What would the results for the process curricula have looked like if only science content had been measured? Their effectiveness would appear to be far less than that which is indicated by the full set of measures. The Follow Through evaluation, unintentionally, may lead to erroneous conclusions by having selected for testing a set of outcomes narrower than the intentions of many of the models it purported to evaluate.

The third impression is that there are curriculum-treatment interactions just as there are aptitude-treatment interactions. Put simply, one instructional method may be best for fostering a particular curriculum outcome but totally ineffective or detrimental for another outcome. Whether or not the second outcome is intentional under the instructional circumstances the effect of instruction on it is there and should not be ignored in a fair evaluation.

The fourth impression is that there must be a better way to assess curriculum impact. The two national efforts I have described cost unreasonable amounts of money given what is likely to be their final impact on American education. To be sure there were many beneficial side efeects of both undertakings, teacher training, heightened interest and concern for children's school experience and so on, but surely there must be easier ways of testing educational models. Already, cheaper, decentralized

curriculum efforts seem to be having some impact and with push from the educational research community they should also be available for shedding light on general curriculum questions.

The final impression is that curriculum can be a powerful societal tool. It does change children and presumably their future lives. Knowing that through the curriculum a society has power to advance basic skills or process or creativity at will, only forces the addressing of more fundamental issues, it does not begin to resolve them.

References

Glass, G.V. Primary, Secondary, and Meta-Analysis of Research. Educational Researcher, 1976, 5, 10, 3-8.

House, E.R., Glass, G.V., McLean, L.D., & Walker, D.F. No Simple Answer: Critique of the Follow Through Evaluation. Harvard Educational Review, 1978, 48, 128-160.

Kennedy, M.M. Findings From the Follow Through Planned Variation Study. Educational Researcher, 1978, 7,6, 3-11.

Stallings, J.A. Implementation and Child Effects of Teaching Practices in Follow Through Classrooms. Monographs of the Society for Research in Child Development, 1975, 40, (2 Serial No. 163).

Stebbins, L.B., St. Pierre, R.G., Proper, E.C., Anderson, R.B., & Cerva, T.R. Education as Experimentation: A Planned Variation Model, Volume IV-A, An Evaluation of Follow Through. Cambridge, Mass.: Abt Associates, Inc., 1977.